Appl. Math. Lett. Vol. 6, No. 2, pp. 55–59, 1993
Printed in Great Britain

0893-9659/93 $6.00 + 0.00
Pergamon Press Ltd

AN ANOMALY CONCERNING TIES IN LOTTO-LIKE GAMES*

GOPALAKRISHNAN SUNDARAM
Department of Applied Mathematics and Statistics
State University of New York, Stony Brook, NY 11794, U.S.A.
STEVEN S. SKIENA
Department of Computer Science
State University of New York, Stony Brook, NY 11794, U.S.A.

(Received October 1992; accepted November 1992)

Abstract—Consider the following game: Each of the n players independently picks a number from 1 through m according to a given probability distribution, and the winner is the player who selected the lowest of the n numbers. A tie occurs if two or more players both selected the lowest number. Surprisingly, contrary to intuition for certain distributions (including one which arises in a version of Bingo, the probability of a tie does not increase montonically. However, as we show the tie probability does increase monotonically when the probability distribution is monotonically non-increasing.

1. INTRODUCTION

Bingo [1] is a variant of lotto. It is a popular game of chance commonly used in North America to raise money for charitable organizations. Each player obtains a card typically in the form of a 5×5 matrix containing 24 distinct integers between 1 and 75, plus a blank as the center entry of the matrix (assume no two cards are identical). From an urn containing balls numbered between 1 and 75, numbers are drawn repeatedly without replacement and are announced. Each player marks the appropriate square on his card if it contains a number drawn from the urn. The winner is the first player whose card is marked in a predesignated pattern, such as filling a complete row or a column. With more than one player in a game, there is a certain non-zero probability that two or more players simultaneously complete the pattern on their cards and that game will end in a tie.

Intuitively, the probability of a tie should increase with the number of players. However, for at least one popular version of Bingo, we shall show that the probability of a tie with two players is higher than the probability of a tie with up to 48 players. We use this anomaly to motivate the study of ties in a more general class of multi-player games.

Consider a generalized game which includes Bingo as a special case. Each of n players selects an integer from 1 to m independently, according to a certain probability distribution (same for every player), where p_i denotes the probability that i will be selected. The winner of the game is the player who selects the smallest of the n numbers. A tie occurs when two or more players select the same lowest number. For example, let $m = 2$, $p_1 = x$, and $p_2 = (1 - x)$, where $0 \leq x \leq 1$. When $n = 2$ a tie occurs only when both the players choose the same number with probability $P_T(2) = x^2 + (1 - x)^2$. When $n = 3$ a tie occurs if all three players choose 1, all three players choose 2, or exactly two players choose 1. The probability of such a tie $P_T(3) = x^3 + (1 - x)^3 + 3x^2(1 - x)$. The anomaly $P_T(3) < P_T(2)$ occurs whenever $x < 1/3$. For $m = 2$, the anomaly is maximized at $x = (8 - 2\sqrt{7})/18 \approx 0.15$, and for this value of x, $P_T(2) = .744$ and $P_T(3) = .674$.

In Section 2, we study the effects of different distributions on this game, and present an algorithm for determining whether there exists an integer n such that $P_T(n + 1) < P_T(n)$ for

*We thank Hussein Badr and Ram P. Srivastav for helpful discussions, and Lisa Sulzberg for instigating this line of investigation.

a particular distribution. In Section 3, we show that monotonically non-increasing distributions never lead to this anomaly. In Section 4, we apply our analysis to Bingo.

The expected number of minima among n d-dimensional vectors drawn from a uniform distribution is analyzed in [2].

2. ANALYSIS

For an implicit probability distribution on m integers, define $P_T(m,n)$ to be the tie probability for an n-player game. A tie occurs on the number i, $1 \le i \le m$ when two or more players select i while the rest of the players pick larger numbers. The probability of a tie on i is given by

$$\sum_{k=2}^{n} \binom{n}{k} p_i^k (p_{i+1} + \cdots + p_m)^{n-k}.$$

Using the binomial theorem, this can be rewritten as

$$(p_i + \cdots + p_m)^n - (p_{i+1} + \cdots + p_m)^n - n\, p_i (p_{i+1} + \cdots + p_m)^{n-1}.$$

Since the tie has to occur with some number, and ties corresponding to different numbers are mutually exclusive,

$$P_T(m,n) = \sum_{i=1}^{m} (p_i + \cdots + p_m)^n - (p_{i+1} + \cdots + p_m)^n - n\, p_i (p_{i+1} + \cdots + p_m)^{n-1}.$$

Since

$$\sum_{i=1}^{m} (p_i + \cdots + p_m)^n = (p_1 + \cdots + p_m)^n + \sum_{i=1}^{m} (p_{i+1} + \cdots + p_m)^n$$

implies

$$\sum_{i=1}^{m} (p_i + \cdots + p_m)^n - (p_{i+1} + \cdots + p_m)^n = 1,$$

we have

$$P_T(n,m) = 1 - \sum_{i=1}^{m} n\, p_i (p_{i+1} + \cdots + p_m)^{n-1}.$$

Define $\Delta P_T(n,m) = P_T(n+1,m) - P_T(n,m)$ as the change in the tie probability when the number of players is increased from n to $n+1$. Obviously,

$$\Delta P_T(n,m) = \sum_{i=1}^{m} n\, p_i (p_{i+1} + \cdots + p_m)^{n-1} - \sum_{i=1}^{m} (n+1)\, p_i (p_{i+1} + \cdots + p_m)^{n}. \tag{1}$$

Using the cumulative distribution function $C_i = p_1 + \cdots + p_i$, equation (1) becomes

$$\Delta P_T(n,m) = \sum_{i=1}^{m} (1 - C_i)^{n-1} p_i ((n+1)\, C_i - 1). \tag{2}$$

In equation (2), the terms will be negative for i such that $(n+1)C_i \le 1$. Observe that $(1-C_i)^{n-1}$ is a decreasing sequence and $(n+1)C_i$ an increasing sequence for $n > 1$.

To test whether a distribution is anomaly-free, it is necessary to provide an upper bound on n such that the first anomaly, if it exists, will occur before n.

THEOREM 1. *For any probability distribution, $\Delta P_T(n,m) \ge 0$ for $n \ge (1-p_1)/p_1$.*

PROOF. When $n \ge (1-p_1)/p_1$, each term in equation (2) will be positive so $\Delta P_T(n,m) \ge 0$. ∎

A particular probability distribution may induce several local minima in the tie probabilities. For example, consider the probability distribution where $p_m = c$ and $p_i = p_{i+1}/r$, $1 \le i \le m$, where r is sufficiently large and c is selected so $\sum_{i=1}^{m} p_i = 1$. When the number of players n is small, all players will select m with high probability. As n increases, with high probability, some player will select $m-1$, decreasing the tie probability. Further increases in n will increase the likelihood of repeated draws of $m-1$, again raising the tie probability. By repeating the argument, it is clear that this distribution will induce $m-1$ distinct minima in the tie probabilities. We leave it as an open problem what is the maximum number of local maxima and minima that the function $P_T(n,m)$ can have.

It is also easy to construct distributions where the position of the first local minima occurs and can be made arbitrarily large, namely that if $p_m > n/(n+1)$,$P_T(k+1,m) < P_T(k,m)$ for all $k \leq n$.

3. MONOTONICALLY NON-INCREASING DISTRIBUTIONS

In each anomalous distribution discussed in the previous section, there is a positive integer i (not exceeding m) such that $p_{i-1} < p_i$. While this condition is not sufficient for an anomaly, it is necessary.

THEOREM 2. *$\Delta P_T(n,m) \geq 0$ for all n, if $p_1 \geq \cdots \geq p_m$.*

PROOF. From equation 2, replacing the second C_i by $p_1 + \ldots + p_i$

$$\begin{aligned}
\Delta P_T(n,m) &= \sum_{i=1}^{m}(1-C_i)^{n-1}p_i((n+1)C_i - 1) \\
&= \sum_{i=1}^{m}(1-C_i)^{n-1} n\, p_i \sum_{k=1}^{i} p_k - \sum_{i=1}^{m}(1-C_i)^n p_i \\
&= \sum_{i=1}^{m}\sum_{k=1}^{i}(1-C_i)^{n-1} n\, p_i\, p_k - \sum_{i=1}^{m}(1-C_i)^n p_i \\
&= \sum_{k=1}^{m}\sum_{i=k}^{m}(1-C_i)^{n-1} n\, p_i\, p_k - \sum_{i=1}^{m}(1-C_i)^n p_i \\
&= \sum_{k=1}^{m} p_k \sum_{i=k}^{m} n\, p_i(1-C_i)^{n-1} - \sum_{k=1}^{m}(1-C_k)^n p_k.
\end{aligned}$$

To complete the proof that $\Delta P_T(n,m) \geq 0$, we prove that for $1 \leq k \leq m$,

$$\sum_{i=k}^{m} np_i(1-C_i)^{n-1} \geq (1-C_k)^n. \tag{3}$$

For $k = m$, both sides of (3) are zero. For $k = m-1$, we have the inequality $np_{m-1}p_m^{n-1} \geq p_m^n$ which is readily satisfied, as $p_{m-1} \geq p_m$ and $n \geq 1$. For the induction hypothesis we assume that (3) is true for $r+1$, $1 \leq r < m$. Now consider the inequality

$$\sum_{i=r}^{m} n\, p_i(1-C_i)^{n-1} \geq (1-C_r)^n. \tag{4}$$

When $n\, p_r \geq (1-C_r)$ then $n\, p_r(1-C_r)^{n-1} \geq (1-C_r)^n$ and hence, equation (4) is satisfied. Thus, it suffices to consider $n\, p_r < (1-C_r)$.

$$\Delta_r = \sum_{i=r}^{m} n\, p_i(1-C_i)^{n-1} - (1-C_r)^n = \sum_{i=r+1}^{m} n\, p_i(1-C_i)^{n-1} - ((1-C_r)^n - np_r(1-C_r)^{n-1}).$$

By the binomial theorem,

$$n\, p_r(1-C_r)^{n-1} - (1-C_r)^n = \sum_{l=2}^{n}\binom{n}{l}(-p_r)^l(1-C_r)^{n-l} - ((1-C_r) - p_r)^n.$$

Hence,

$$\begin{aligned}\Delta_r &= \sum_{i=r+1}^{m} n\,p_i(1-C_i)^{n-1} - (((1-C_r)-p_r)^n - \sum_{l=2}^{m}\binom{n}{l}(-p_r)^l(1-C_r)^{n-l}) \\ &= \sum_{i=r+1}^{m} n\,p_i(1-C_i)^{n-1} - ((1-C_{r+1})+p_{r+1}-p_r)^n + \sum_{l=2}^{m}\binom{n}{l}(-p_r)^l(1-C_r)^{n-l}.\end{aligned}$$

By the induction hypothesis

$$\sum_{i=r+1}^{m} n\,p_i(1-C_i)^{n-1} \geq (1-C_{r+1})^n \geq ((1-C_{r+1})+p_{r+1}-p_r)^n$$

as $p_r \geq p_{r+1}$. It remains to show that $\sum_{l=2}^{n}\binom{n}{l}(-p_r)^l(1-C_r)^{n-l} \geq 0$. This sum alternates between positive and and negative terms and starts with a positive term. Consider the difference Δ_l between any two successive terms, the first being positive.

$$\begin{aligned}\Delta_l &= \binom{n}{l}p_r^l(1-C_r)^{n-l} - \binom{n}{l+1}p_r^{l+1}(1-C_r)^{n-l-1} \\ &= p_r^l(1-C_r)^{n-l-1}\left(\binom{n}{l}(1-C_r) - \binom{n}{l+1}p_r\right) \\ &= p_r^l(1-C_r)^{n-l-1}\left(\binom{n}{l}(1-C_r) - \frac{n}{l+1}\binom{n-1}{l}p_r\right).\end{aligned}$$

We know that $\binom{n}{l} \geq \binom{n-1}{l}$ since $\binom{n}{l} - \binom{n-1}{l} = \binom{n-1}{l-1}$. By our assumption, $(1-C_r) > np_r$ and hence, $(1-C_r) \geq np_r/(l+1)$ as $l > 1$. Thus,

$$\left(\binom{n}{l}(1-C_r) - \frac{n}{l+1}\binom{n-1}{l}P_r\right) \geq 0,$$

which implies that

$$\sum_{l=2}^{n}\binom{n}{l}(-p_r)^l(1-C_r)^{n-l} \geq 0$$

and the proof is complete. ∎

4. BINGO

Bingo can be considered as a special case of our generalized game by the following argument. We limit discussion to the version of Bingo where the winner is the first player to completely cover their card. Instead of using repeated draws from the urn to define a permutation on the possible card values, assume the identity permutation is drawn and have each player select uniformly at random a card containing s entries between 1 and m. The winner is the player whose maximum number is minimized. Since there are are $\binom{m}{s}$ possible cards, $\binom{i-1}{s-1}$ of which contain a maximum value of i, the probability distribution is given by

$$p_i = \binom{i-1}{s-1}\Big/\binom{m}{s}.$$

This probability distribution monotonically increases with i, since $\binom{i-1}{s-1}$ is an increasing function of i. Using equation (2), the tie probability is given by

$$P_T(n,m) = 1 - \sum_{i=1}^{m} n\left(\binom{i-1}{s-1}\Big/\binom{m}{s}\right)\left(1-\binom{i}{s}\Big/\binom{m}{s}\right)^{n-1}.$$

Table 1. $P_T(n, 75)$ for full cover Bingo.

n	$P_T(n,75)$	n	$P_T(n,75)$	n	$P_T(n,75)$	n	$P_T(n,75)$
2	0.19313	3	0.17604	4	0.17637	5	0.17806
6	0.17930	7	0.18026	8	0.18108	9	0.18182
10	0.18250	11	0.18311	12	0.18368	13	0.18420
14	0.18469	15	0.18515	16	0.18557	17	0.18598
18	0.18636	19	0.18672	20	0.18707	21	0.18740
22	0.18771	23	0.18801	24	0.18830	25	0.18858
26	0.18885	27	0.18910	28	0.18935	29	0.18959
30	0.18983	31	0.19005	32	0.19027	33	0.19048
34	0.19069	35	0.19089	36	0.19109	37	0.19128
38	0.19146	39	0.19164	40	0.19182	41	0.19199
42	0.19216	43	0.19232	44	0.19248	45	0.19264
46	0.19279	47	0.19295	48	0.19309	49	0.19324

Numerical results for $P_T(n, m)$ when $m = 75$ and $s = 24$ appear in Table 1. Table 1 illustrates that $P_T(n, m)$ grows monotonically for $n \geq 3$, but slowly enough that ties with up to 48 players are less common than a tie with two players.

REFERENCES

1. J. Scarne, *Scarne's Complete Guide to Gambling*, Simon and Schuster, New York, (1961).
2. M. Schkolnick, J.L. Bentley, H.T. Kung and C.D. Thompson, On the average number of maxima in a set of vectors and applications, *J. ACM* **25**, 536–543 (1978).